DOCUMENT RESUME

ED 027 073 PS 001 549

By-Baumrind, Diana

Naturalistic Observation in the Study of Parent-Child Interaction.

California Univ., Berkeley. Dept. of Psychology.

Pub Date 68

Note-20p.; Paper presented at symposium at the 76th American Psychological Association convention, San Francisco, California, August 30 to September 3, 1968.

EDRS Price MF-$0.25 HC-$1.10

Descriptors-Behavior Rating Scales, Child Rearing, Family Environment, Item Analysis, Laboratory Experiments, *Observation, Parent Attitudes, *Parent Child Relationship, Parent Influence, *Preschool Children, Research Methodology, *Research Problems, *Social Development, Socialization, Test Reliability, Theories

Identifiers-Parent Attitude Inquiry, Parent Rating Scales, Preschool Behavior Q Sort, Stanford Binet

This project investigated patterns of parental authority among Berkeley preschool children and the processes by which these parents contributed to the development of children's social responsibility and individuality. Subjects were 140 families from city-sponsored, private cooperative, and university-operated nursery schools. Eight constructs were devised: (1) high vs. low stress tolerance, (2) self-confident vs. fearful, (3) achievement-oriented vs. nonachievement-oriented, (4) approach-oriented vs. withdrawn, (5) autonomous vs. suggestible, (6) rebellious vs. dependable, (7) destructive vs. constructive, and (8) alienated vs. trusting. Observation data on parent behavior were rated on scales approximating child scales and self-report parental attitudes were collected. Research is incomplete, but laboratory experimental and naturalistic observations were assessed and the latter was favored. It was concluded that observational studies which focus on the human psyche and human behavior seldom can achieve situational control, reliability of measurement, or precise formulation of process variables. They can, however, proceed self-critically, using statistical tests of significance on well-formulated hypotheses which are well defined conceptually and operationally. (DO)

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

NATURALISTIC OBSERVATION IN THE STUDY OF PARENT-CHILD INTERACTION

Diana Baumrind, Department of Psychology, University of California, Berkeley

Presently I am investigating patterns of parental authority among parents of Berkeley preschool children, and the processes by which these parents may contribute to the development in young children of social responsibility, and individuality. This study is an extension of my previous work, already reported, on socialization practices associated with dimensions of competence in preschool boys and girls. The present study differs from the previous study in that the sample is more broadly representative of Berkeley parents; new variables reflecting current ideology and practices have been added; and new measures, including a self-report inquiry, have been developed.

It is clear that the social revolution now in progress has fundamentally altered the attitudes of youth towards established authority as well as the attitudes of those who have authority. Many of the parents participating in the present study have rejected traditional values, and as part of their rejection of many aspects of the social structure, are searching for ways of redefining the parent-child relationship. The parents are also skeptical of the relevance to their problems of social science research. As a consequence, the role of the research scientist in relation to his subjects will, I think, have to undergo the same radical change as the role of the mental health professional in relation to his patients. This loss of faith in social institutions is clearly apparent to the observer working in natural settings.

Outline of the Study

Now to the study itself. I am studying about 140 families drawn from city-sponsored, private cooperative and University-run nursery schools in Berkeley.

Data on Child Behavior

Data on child behavior were collected during three to five months of intensive observation of the child in the nursery school setting and while taking the Stanford-Binet. Observation over a 3-5 month period permits the observer to distinguish

---

This paper was prepared for the 76th American Psychological Association convention symposium: Approaches to and applications of the study of parent-child interaction. San Francisco, August 30 to September 3, 1968.

ED001549
PS001549

enduring response tendencies from maturational factors. As Dr. Heinicke has pointed out, repeated samples of behavior are needed in order for the observer to note how the child reacts to a wide range of stimulus objects and external cues. The child's emotional lability and his responsiveness to temporary situational factors precludes the possibility of evaluating a young child after a few brief observations.

Each child in the study was rated on a 72-item Q sort which defines operationally, for the nursery school setting, 8 constructs. The constructs are designated as: 1. High vs. Low Stress Tolerance; 2. Self-confident vs. Fearful; 3. Achievement-oriented vs. Nonachievement-oriented; 4. Approach-oriented vs. Withdrawn; 5. Autonomous vs. suggestible; 6. Rebellious vs. Dependable with Adults; 7. Destructive vs. Constructive; 8. Alienated vs. Trusting. Nine items were devised to explore each construct. An item was defined by describing: how a child rated high and rated low would behave in the nursery school setting; pertinent situations in which the child would be likely to demonstrate the behavior; and how each item differed from other items intended to measure similar kinds of behavior. For example, Item 1:

Item 1. Expresses negative feelings openly and directly (frustrated or bound up by feelings)

High: The child does not become bound up with negative feelings but is able to express himself openly and with appropriate emotionality relevant to the cause and cure of what is troubling him when experiencing a generally frustrating situation, particularly when an altercation arises with a peer or a divergence with a teacher, e.g., he cries to get help or he expresses anger, openly and directly.

Low: In a frustrating situation the child is likely to become bound up with feelings and become either excessively passive or regressive, e.g., he cries because of inability to express anger or to take action.

Pertinent Situations and Behaviors: Note reactive regressive, or passive behavior under stress particularly during divergences with peers and teachers. Note when teacher asks child to give up playing if he becomes angry or disappointed but cannot express it and becomes resentful.

> Differentiations: This item should be differentiated from Item 29, Emotionally expressive (emotionally bland), which focuses on the ability or willingness of the child to express positive emotions, although children typified by bland affect are apt to receive low ratings on both items. (Baumrind, *Manual*, 1968)

The observer staff, meeting jointly, criticized the definitions of the items and in the process constructed a manual for the use of the Q sort.

Data on Parent Behavior

Data on parent behavior were collected by observers in the home with the entire family present, during the last 3 hours preceeding the child's bedtime. Each parent was interviewed separately, and then rated. The items on the Parent Rating Scales were constructed and defined in much the same way as the items on the *Preschool Behavior Q Sort*. Five items operationally define each of 15 constructs, such as: I. Consistent articulated childrearing philosophy vs. lack of philosophy; II. Directive vs. non-directive behavior; III. Firm vs. lax enforcement policy; IV. Confidence vs. lack of confidence in self as a parent; V. Discourages vs. encourages infantile behavior; and so on. For each item, alternatives were designated 1, 1+, 2, or 2+. The observer first decided which of two forced-choice alternatives best described the parent and then the degree to which the parent was characterized by the behavior rated. The + alternatives, which signify a more intense degree of 1 or 2 were also defined. 0 was used as an alternative when a choice between the two specified alternatives could not be made because of lack of information, or because both alternatives assume something which is not true for that particular family. Some items have 0-1 & 0-2 specified, as will be shown in the example to be cited. Where an item was particularly apt as a description of a parent, its number was circled. Where an item was irrelevant or the observer regarded the rating as unreliable, the number was slashed. In addition to a

separate rating of the mother and the father, the family as a whole was rated globally on each of the 15 Childrearing Dimensions, defined conceptually.

One item intended to measure the construct, Firm vs. lax enforcement policy, is entitled *Cannot be coerced by child.* Two meanings of 0 are specified which would preclude use of the item to measure the construct, Firm vs. lax enforcement policy.

0-1 Child never uses coercive tactics

0-2 On principle, parent does not oppose child's will

Alternative 1 reads:

> Child can overcome parental oppositions by crying or causing a commotion (+ In addition, parent is frequently coerced into complying with child's wishes by his use of such tactics)

Alternative 2 reads:

> Child cannot overcome parental opposition by crying or causing a commotion, although he may succed by use of reason (+ Parent firmly opposes child when he uses such tactics and never succembs to child's nuisance value)

Parent Attitude Inquiry

Self-report data on parental attitudes were obtained from each parent prior to the home visit. The construction of the Parent Attitude Inquiry paralleled that of the other measures in that it included opportunity for feedback from the user, in this case the parents. In the initial versions of the inquiry, a number of situations were described which might arise between parent and child. Each situation was followed by several Likert-type statements in which the respondent indicated on a 5-point scale, from 'strongly agree' to 'strongly disagree', the extent to which he or she supported an action which the parent in the constructed story might take. The situations were drawn from actual encounters described by parents in interviews in the previous study and were quite realistic. However, the initial format was abandoned because too many of the 200 parents with whom the inquiry was discussed, objected to the encounters described and the alternatives offered as not applicable to their particular situation.

The fact that the situations described, while realistic, were not always meaningfully related to a given family's situation proved unexpectedly annoying and distracting to many of the parents. I learned about their feelings by providing room for comments connected with each situation described, and by talking to 12 groups of about 25 parents each. In response to the parents' objections and the content of their suggestions, I then devised a format using forced-choice questions, similar to the format used by Loevinger and Sweet on the Family Problem Scale (1961). Not many of the actual items in the Loevinger and Sweet scale could be used because Berkeley parents found them too stereotyped, out of date, and oversimplified, and felt they could identify good and bad poles. Many of the original items I devised had to be eliminated after the pretest, for the same reasons. The second version was revised after being administered to another 200 parents who commented in writing on the items. The present version consists of 117 items. One-third of each page is left blank for comments. Also for each item, the respondent can, after choosing one of the two alternatives offered, check a column indicating that neither alternative really applies. The spaces left for comments were used extensively by parents to criticize the items and clarify the objective responses. Subjects were critical of unreal alternatives, overstated alternatives, alternatives which were not mutually exclusive and so on. Their comments reflect fundamentally their concern that the investigator will stereotype rather than understand them. I don't yet know whether the Parent Attitude Inquiry predicts parent or child behavior since data collection is still in progress. However, based on parents' reactions to the inquiry, I have serious doubts about the possibility of getting rid of what is called response set in a self-report instrument. What is called response set seems to reflect among other problem-solving maneuvers the respondent's way of coping with a double-bind communication. One message to the parent is that he is to tell the investigator what he or she really thinks or believes. The conflicting message is that he is to do so only within the confines of the alternatives offered. Where these messages

PS 001549

are incompatible, the respondent cannot make a rational decision and becomes justly critical of the item. The respondent finds that the alternatives offered are too obvious; or, that the alternatives are not obvious enough which then leads him to search for the real meaning behind the question. The respondent may agree with the sense of the question but disagree with the qualifiers. The respondent may be unable to affirm an alternative with which he agrees because it is too boldly stated or because it implies a negative or a positive evaluation of another. He may instead feel forced to choose the alternative with which he agrees less but whose wording is more acceptable. To the extent that there are many different types of response sets operating simultaneously, and to varying degrees in different subjects, such response tendencies cannot be controlled by a choice of format or control of wording.

After this experience, I much prefer an interview to a self-report inquiry as a method for assessing parental attitudes and values. My assumption that parents would prefer to take a self-report inquiry in the comfort of their own home than to be interviewed was incorrect. The parents we studied preferred for the most part to discuss those views which mattered to them, in depth, in an interview. So involved in the issues discussed were most parents that many of the interviews, although highly structured, consumed over 25 single-spaced pages when typed. An interview is also an essential adjunct to observation since the observer can then investigate the internal instigators of the parents' behavior.

Despite the title of this paper in the program, I intend to focus most of my discussion on naturalistic assessment of parent-child interaction, and the reasons why I find assessment in the home setting valuable despite its many serious methodological drawbacks.

Laboratory Experimental Contrasted with Naturalistic Observation

Research studies of parent-child interaction are justified ultimately by their usefulness to those who rear children and to the experts who advise them. All procedures

used to study a phenomenon change that phenomenon. An animal does not react in a laboratory in the same way as he does in his natural habitat. The more the investigator controls the research environment the more likely he is to be able to take reliable measurements on his animal, and to make cause and effect statements concerning the determinants of the animal's behavior in the laboratory. Laboratory-experimental procedures are designed to contain and control human behavior so that causal relations between antecedent and consequent conditions can be established. The possibility of generalizing from behavior observed in the laboratory setting to behavior observed in the natural setting is in part a function of the changes which containment and control produce in that behavior. If such containment and control radically alter the effect of adult behavior on the child, then generalizations of findings from the laboratory to the home setting are invalid.

For example, let us examine studies which investigate effects of variations in timing, or in severity of punishment on response inhibition. Suppose that these experiments are performed on adolescent rats. To what extent can these findings be pertinent to adolescent humans? If a rat, shocked within 1/5 of a second, learns to avoid a blind alley better than when shocked 5 seconds after he has made the wrong move, will an adolescent administered a sharp slap within 1/5 of a second after uttering an obscene remark to his mother be less likely to repeat his offense than if slapped after 5 seconds? If the intensity of electric shock affects the rat's subsequent responses, will the hardness of the slap affect the human adolescent's subsequent responses? Clearly, it is impossible to defend on rigorous grounds such generalizations. First, the parameters designated as equivalent must be operationally defined on empirical grounds, by observation of the adolescent in natural settings. Second, the differential effects introduced by species-specific and age-specific factors must be considered. For example, an adolescent tends to rebel against an arbitrary act of authority. Extremely early punishment prevents use of reason and so may be construed

as arbitrary. For species and ages where use of reasoning induces compliance, 'early' punishment by preventing use of reason may provoke rather than deter the proscribed response.

The actual times which define 'early' and 'late' punishment when a differential effect on response inhibition could be demonstrated in the laboratory cannot be replicated meaningfully in the home. Parke and Walters (1967) in order to replicate Walters' own work in respect to timing of punishment effects, were forced to define 'early' as prior to the time the proscribed act was complete; and 'late' as 5 seconds after the proscribed act was completed. Even a parent with rapid reflexes cannot react <u>that</u> rapidly in the home setting.

Let us take another example of the way in which the situational differences restrict generalizability of results. Let us examine the effects of level of nurturance on imitative behavior, as this effect interacts with the length of time the child knows the experimenter. Bandura and Huston (1961) exposed one group of children to a female model who was highly nurturant and rewarding, and a second group of children to a nonnurturant model who was distant and nonrewarding. The group of children exposed to the nurturant model imitated the model's responses (other than aggression) more than the group of children exposed to the nonnurturant model. Madsen (1968) repeated this experiment in most of its details, except that the child was exposed to the two models over a 6-week school session. He found no differences in imitation of the model's responses. If Madsen's results are accepted as reliable, then the likelihood of specific imitation of adult behavior seems to be reduced rather than enhanced by repeated exposure, at least under the circumstances cited. Moreover, in the toy-rejection condition the effect of a model who rejects a particular toy decreases significantly when the child has prior experience with the toy involved. The longer the period of time the child knows the experimenter and the more familiar he is with the toy involved, the more the experimental condition resembles that of the home.

Conditions in the laboratory differ systematically from conditions in the home in several ways:

1. Unfamiliarity with the adult, the setting, *and* the stimulus objects characterizes the experimental situation, while familiarity with the adult, the setting, and the stimulus objects characterizes the home situation; familiarity with the adult, the setting, and the stimulus objects should interact with most independent variables in their effects on child behavior, as was demonstrated by Madsen's attempted replication of Bandura's work, just cited. When the parent is used as the experimenter, the stimulus objects and the setting are unfamiliar, and the parent's behavior may also be construed as odd by the child.

2. In the laboratory setting, the adult, even if the adult is the parent, has maximum control over his own behavior *and* can program that behavior without much regard for the reactions of the child-subject. The E is sure of himself; he operates in accord with principles *and* values which he has fully integrated; *and* he often derives his sense of purpose *and* rectitude from a higher authority, the principle investigator. His relationship to the child has much in common with the authoritarian parent of a previous era who acted with unambivalence conferred by tradition *and* theistic convictions. The self-assured, single-minded adult experimenter has little in common with the modern parent. Even the parent acting as experimenter assumes conferred authority which the same parent may lack in the home. A model should be more effective when he acts in a clear, self-assured manner. He should be more effective when the child he is attempting to influence is unclear *and* lacking in self-assuredness, as indeed the child is in the laboratory, due to his lack of experience with the stimulus objects or the rules of the game.

3. Thirdly, reciprocity of reinforcement characterizes normal parent-child interaction. By virtue of the control which E must have over his own behavior, the E-child interaction lacks reciprocity. While E cannot force the child to obey in the same way a parent can, he is set to resist the negative sanctions *and* positive reinforcers applied by the child, to which normal parents respond spontaneously.

Patterson and Reid (1967), among others, indicate that nonreciprocity of reinforcement characterizes the relationship of neurotic children and their parents. When the parent acts as an E, his construction of what the investigator expects probably reduces his normal responsiveness to the child's behavior. In this sense, the laboratory-experimental situation simulates one kind of neurotic parent-child relationship, the kind in which the parent is not responsive to social reinforcement by the child.

4. Fourth, the child's responsiveness to the adult in the laboratory is effected by different factors than in the home. Techniques which induce behavior change in the laboratory where the child is striving to learn the rules of the game and thus is responsive to reinforcement may not do so in the home where the rules of the game both are known to the child, and manipulatable by him. Under other circumstances the child develops a negative set and resists experimental directives of a kind he would accept from adults he loves. Patterson (1967) among others suggests that responsivesness of a normal child to an adult model varies as a function of a complex interaction between sex of both the child and the model, and the age of the child. If that is so, the responsiveness of the child to a male E should in many instances fail to predict his responsiveness to his mother in the home. To the extent that responsiveness is a necessary precondition to behavior control achieved through imitation or reinforcement, the results of studies on behavior control in the laboratory will not generalize to behavior control in the home.

Naturalistic observation is used because conditions in the laboratory differ systematically, in the ways discussed, from conditions in the home. It is also true that naturalistic observation vicariously involves the investigator in social phenomena permitting him to develop hypotheses consistent with the changes in the phenomena studied. For example, the component attitudes which define contrasting types of parents are rapidly changing. Social responsibility is not defined by obedience and hard work in the way it was five years ago. Expressive qualities are much more highly valued today even by rather rigid, controlling parents.

Limitations/Advantages of Naturalistic Observation

We have discussed some of the characteristics of laboratory-experimental methods which from one perspective limit their usefulness. Naturalistic observation possesses the complementary set of limitations/advantages. The choice between these extremes in methods and among all the many variations of these methods should be determined by the actual phenomena to which interpretations are meant to apply. Only when one method has been shown to be superior for a given purpose should another method be discarded for that purpose. For example, it appears that many parents are not accurate informants about their nonspecific childrearing practices unless they have been observed in action with their children so that nonspecific childrearing practices should probably be observed rather than reported; and that childrearing values are better assessed by an interview than by a self-report inquiry. The scientific superego of the investigator should dictate not the choice of method as much as the standards relevant to a given method.

Lack of Control

1. Laboratory-experimental methods permit the investigator to control the adult behavior whose effects upon the child are being studied. Naturalistic observation precludes such control. While laboratory methods require that the subject meet the E on the latter's home ground, observational procedures require that the *observer* meet the S on his home ground--literally, in the study of parent-child interaction. Conditions can be standardized somewhat, but within broad limits must be allowed to vary freely. Control as achieved in the laboratory-experimental setting is neither desirable nor possible is most studies of normal parent-child interaction in the home. The investigator does not want to alter the pattern of behavior he is observing and unless a parent has come for help, the parent does not want to have his behavior manipulated in his own home.

For example, the conditions of observation in our study were standardized in

that each family was observed twice for 3 hours, roughly between the hours of 5-8 p.m., and all family members were required to be present. The observer wrote continuously and did not interact. By these conditions we sought to minimize facade and to elicit a high rate of theoretically relevant behavior. However, controls such as these would not provoke, for example, sufficient instances of punishment so that we could study timing or severity of punishment in any systematic manner. Without controls which distort the data, we may not be able to observe the phenomenon systematically enough to study it at all. Perhaps we could train a family member to act as an observer or to introspect systematically if we wanted to obtain enough instances of punishment but these techniques also have their own serious drawbacks.

Unreliability of Measurement

2. Another serious methodological problem inherent in naturalistic observation is unreliability of measurement, as a function of variations among observers, and among samples of parent behavior. Efforts to increase reliability are often impractical or reduce the validity or relevance of the data collected. In theory, reliability of ratings can be greatly increased by having many observers take repeated measurements. Jack Block (1961), among other methodologists, has observed that the best solution to the problem of observer agreement is to composite judgments from three or more raters. However, in practice, budgetary limitations prohibit the use of multiple observers on a reasonably large sample of families. In addition, the presence of multiple observers would disrupt normal household routine more than the presence of a single observer. The number of visits which can be made to a given home is also limited, in practice, by cost, and the subject's tolerance of the invasion of his privacy.

While families coming to a clinic for help are often willing to tolerate repeated visits so that they can be helped, families who do not seek or expect to receive direct

personal help are less acceptant of repeated visits. I do believe, however, that at least two visits to each home are necessary both to obtain reasonably stable samples of behavior, and because the observer and parents are able to behave more spontaneously during the second visit.

Reliability can be increased by discouraging dissimulation. Between the hours of 5 and 8, when our visits were made, parents of preschool children have so much to cope with that dissimulation is difficult. The evocative behavior of young children and the many tasks which must be accomplished at that time of evening reduce facade effect. Any marked dissimulation by the parents is revealed in the actions and comments of the children. Of course, the parent generally extends the best side of himself, and inhibits what he regards as outrageous behavior. As a consequence, we do not witness many episodes of rage or intimate fondling, although we do witness some.

The collection of data from several sources should enhance reliability of rater judgments. In addition to being observed, each parent was interviewed about his or her reactions to the home visits. Parents acted as informants concerning their intentions in a sequence when these were not apparent, and they added relevant behavior which the observer may not have witnessed. Data provided by parents concerning their values, their typical behavior during other times of day, and the child's behavior at an earlier age were probably more valid and less effected by wishful thinking and fantasy than would have been the case if the observer had not been a witness to actual parent-child behavior.

The observer is in a better position to record reliably relevant events if he is thoroughly familiar with the rating or coding categories which he will later use to quantify the data he is collecting. It is important that members of the observer staff have an objective and shared basis for deciding what events take priority over other events since they can never record all that occurs in the home.

I do not believe that reliability of measurement should be achieved by selecting raters who share the hypotheses of the investigator, or who share with each other

similar viewpoints. By selecting, instead, observers representing a wide diversity of training, personal styles, and ideological convictions, the fine points of the rating system can be hammered out and disagreements due to misunderstandings and ideosyncratic prejudices resolved by the process of critical group interaction.

An additional technique which may increase reliability is to have observers monitor their own ratings for response tendency. By recording their own scores, case by case, observers can compare scores across cases and in that way make judgments which are more truly normative.

Naturalistic observation is most meaningful when children are under 6 and when observers and the families they visit share a common cultural heritage. Older children do not interact at the same rate of intensity with their parents, and are probably more effected by the presence of an observer even than their parents. For families of older children, and where the observer would be felt as an intruder, as is true when a white observer enters a black family's home, the use of miniature situations seems to be one excellent alternative to naturalistic observation.

Clearly then, naturalistic observational methods possess serious limitations which to some extent are inherent in the methods. This is true of all scientific methods, especially when applied to psychological phenomena. If subjected to the same type of scrutiny, the developmental method used by Kohlberg to study parent-child relations would be shown to be as fallible as the laboratory-experimental or observation methods, although its limitations are different. Thus, the developmental method assumes a basic and powerful universal urge in human beings toward higher levels of development. But the relative absence of this urge in many people is itself a problem much in need of study. The developmental method assumes that the organism finds interesting that which is somewhat difficult for him. Yet this too is a dimension along which people differ. Developmental laws are general-type propositions and as such apply equally to each and every member of the group. For general-type propositions, the exception does indeed probe the rule. Yet few serious efforts are made to explain the large percentage

of exceptions, e.g., in the supposedly invariant order of progression in levels of moral judgment when actually studied empirically.

The Place of Theory in Observational Research

I would now like to discuss some of the sources of variations in different investigators approaches to naturalistic methodology. Until recently, observational research tended to be atheoretical on philosophical grounds. In discussing the type of protocol which the observer should collect, Barker defined a specimen record as

> a sequential account of a long segment of a person's behavior and situation as seen and described by a skilled observer. It reports in concrete detail a stream of behavior and psychological habitat. Specimen records do not interpret behavior within the framework of psychological theories; they describe behavior in the concepts and language of laymen; they provide unanalyzed, theoretically neutral data that can be used for many different purposes (1961, p. vi).

By contrast, I regard naturalistic observation as most meaningful when it is guided by theory at all stages. My use of observational methods is more in accord with Weick's broader definition of observation methods as "the selection, provocation, and recording of that set of behaviors, settings, and codes concerning people *in situ*, that is consistent with empirical aims" (in press). It is the function of theory to select relevant categories in accord with the problem as defined by the investigator. Human behavior, including that of the investigator is purposive, and the partial narrative record of family interaction should provide data relevant to the research purposes of the investigator. The observer should know what categories will be used to code the behavior he is recording so that his record can cover as completely as possible those areas of behavior of greatest interest to the investigator.

Issue of Natural Units

I have discussed elsewhere (Baumrind, 1967) the interaction unit which we used to code the home visit transcript. I will not describe here how we define a

sequence but rather discuss one issue which arises in defining a unit, i.e., the rather curious issue concerning the relative merits of natural versus artificial units. I say 'curious' issue because I cannot imagine how a unit could be natural. In order to understand human behavior, the investigator must superimpose structure upon the ongoing processes observed. In effect, he stops the ongoing stream of behavior by whatever techniques of recording and quantification he uses. The very act of observation requires that he record in static unambiguous signals what are in fact ambiguous, ever-changing phenomena. Certainly the stream of behavior cannot *occur* in perceptible units. Rather, the stream of behavior must be *organized* into units by the observer whose unique apparatus, of necessity, alters the reality out there. A process cannot consist of natural units, because units are not natural to a process. A process is infinitely divisible.

Units are requisite to the *understanding* of a process, however. The investigator defines his own unit in relation to that aspect of the process selected by him for special attention. Once the investigator has made explicit what aspect of that process is relevant to his objectives, that is, once he has decided what it is he wants to know, he can then define *his* unit of analysis and the class of events to which that unit corresponds. When perceived and quantified, the event of necessity is altered because it is removed from context. But the event must be removed from the larger context to be studied, even when some aspects of the immediate context are retained. Since units are unnatural to process, it seems pointless to distinguish among units on the basis of whether they are natural or artificial. A 'good' unit is meaningfully related to the investigator's problem and able to be identified reliably.

Observer Inference and Intuition

Some investigators would like to exclude intent as a basis for coding the actor's behavior in an interaction. However, in order to rate parent-child interaction or

identify a sequence of interaction, it is necessary for the observer to *infer* the goals of the actor from his actions or where necessary from a later interview. In rating participants, the human observer is capable of discounting acts and statements which are produced by a fleeting mood or set in favor of more permanent essential components of human interaction, and this capability can increase either the truth or error component of his judgment.

As Tinbergen (1968) points out, all animals are skilled at interpreting accurately nonverbal signals and gestures. This sign language, which is species-specific and frequently cross-cultural reflects our true motives better than speech and is probably the basis for what we call empathic understanding in human beings. The investigator concerned with purposive behavior must be concerned with motives. Observers should be trained not to suppress empathically-achieved knowledge but rather to make explicit those cues which allow them to know empathically what is going on. Knowledge based on empathy is probably at least as reliable--if reliability is defined as observer agreement--as knowledge based on induction. It is important to permit observers to credit intuitive judgment, but then to insist that they justify their intuitive perceptions by specifying the behavior used to rate the participants.

The apparent objectivity which the laboratory-experimental method possesses can not be appended to the observation methods by excluding intent and focussing on effect in order to define an interactive unit. While it is impossible for the observer to ignore the effect of a parental action on a child, he cannot use the child's response as the major criterion for coding a parental act without hopelessly confounding independent and dependent variables. Neither can contiguity of two acts be used as the sign that they are stimulus-response connected. Contiguity is an unreliable cue to the identification of causal connections between two physical or verbal acts of human beings, since delayed responses characterize organisms capable of abstraction.

Quantification of Observational Data

Investigators disagree on the way in which observational data should be handled. Observational methods are seductively gratifying to use because they permit the observer to confront reality at the point of data collection. The observer may have the impression that he 'knows' his phenomena directly, in contrast to 'knowing about' these phenomena. However, the scientific or lay reader cannot possibly duplicate the experience of the observer and therefore must have a basis for *evaluating* the accuracy and the generalizability of what the observer 'knows'. Often the investigator is acutely aware that in the process of abstracting, quantifying, and systematically summarizing his data, much of the concrete truth which the observer 'knows' is lost and distorted. He wishes, therefore, to use a natural history approach, and to emphasize idiographic rather than nomothetic events. The reader of a scientific report, however, is limited to 'knowing about' the phenomena which the observer may feel he 'knows'. The reader cannot rationally be expected to share the observer's inner certitude achieved via personal encounter. Unlike some other sciences which do make productive use of natural history methods, the data of psychology (especially psychological interaction) are so infinitely varied and varying,and intrinsically endowed with subjective meaning that the natural history approach of presenting raw data is of limited usefulness. Naturalistic assessment then should be differentiated from natural history studies. The former may have usefulness in the study of parent-child interaction far beyond that of the latter with which it is sometimes confused.

Conclusion

I would like in closing to make some general remarks on method in psychological research, and to relate these remarks to the use of naturalistic observation as a method.

Psychology has as its domain, the human psyche. There is no domain of science more difficult to master than the human psyche. This is so for many reasons:

1. The human mind is both the source and the object of knowledge. Subjectivity in psychological research can be regulated but not overcome.

2. Human subjects do not want to be known. Like investigators, they may be willing to predict and control the behavior of others, but they are not willing to be predicted and controlled themselves.

3. Psychological phenomena change very rapidly. Man alters himself in response to his own ideation. Freudian man stopped being a reality almost as soon as he learned that he existed. The same is now happening to Skinnerian man. Man uses the knowledge which he possesses concerning the forces which control him, to neutralize those forces.

The purpose of scientific methods is to make possible additions to the collective knowledge of man. When the psychologist dons his scientific hat, he agrees to abide by explicit rules of evidence, which vary somewhat with the method he has adopted. He accepts these constraints in order that he can persuade other members of the scientific community of the truth of an idea by means both accept as binding. Scientific standards diverse as they may be have a common aim, to permit the investigator to evaluate the probability that a proposition believed to be true, is false. Only by adhering to explicit standards associated with a given method can a scientist add to collective knowledge.

The philosophy and techniques of naturalistic observation permit the investigator to focus upon the proper domain of psychology, the human psyche, and to study internal instigators of behavior by introspective, empathic, and inferential processes. Observational studies which focus upon the human psyche as well as human behavior seldom can achieve situational control, reliability of measurement, or precise formulation of process variables. Such studies can, however, proceed self-consciously and self-critically, using statistical tests of significance on well-formulated hypotheses, and constructs which are well defined conceptually and operationally. In my opinion, naturalistic observation does not add to the collective knowledge of man unless it meets those standards. The subjectivism inherent in observational methods must be reduced by heterodox views within the staff, explicit formulation of testable hypotheses, and the requirement that observers justify their intuitive, empathic, and inferential judgments by specifying the behaviors used in arriving at these judgments.

References

Bandura, A., & Huston, A. C. Identification as a process of incidental learning. *Journal of Abnormal and Social Psychology*, 1961, *63*, 311-318.

Barker, R. G., Barker, L. S., Schoggen, M., & Wright, H. F. *Specimen records of American and English children*. Lawrence, Kansas: University of Kansas, Publ., Soc. Sci. Stud., 1961.

Baumrind, D. Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*, 1967, *75*, 43-88.

Block, J. *The Q-sort method in personality assessment and psychiatric research*. Springfield, Ill.: Charles C. Thomas, 1961.

Loevinger, J., & Sweet, B., Construction of a test of mothers' attitudes. In J.C. Glidewell (Ed.) *Parental attitudes and child behavior*. Springfield, Ill.: Charles C. Thomas, 1961. Pp. 110-123.

Madsen, C. Jr. Nurturance and modeling in preschoolers. *Child Development*, March 1968, *39-1*, 221-236.

Parke, R. D., & Walters, R. H. Some factors influencing the efficacy of punishment training for inducing response inhibition. *Monograph, Child Development*, 1967, *32-1*, 1-45.

Patterson, G. R., Social learning: An additional base for developing behavior modification technologies. In C. Frank (Ed.) *Assessment and status of the behavior therapies and associated developments*. New York: McGraw-Hill, 1967.

Patterson, G. R., & Reid, J. Reciprocity and coercion: Two facets of social systems. Paper prepared for the Ninth Annual, Institute for Research in Clinical Psychology, sponsored by the University of Kansas, Department of Psychology: Behavior modification for clinical psychologists. Lawrence, Kansas. April 1967.

Tinbergen, N. On war and peace in animals and man. *Science*, 28 June 1968, *160-3835*, 1411-1418.

Weick, K. E. Systemic observational methods. In Lindzey and Aronson (Eds.) *Handbook of Social Psychology*. Reading, Mass.: Addison-Wesley, 1968, in press.